Run 3 PHENIX DAQ Overview

John Haggerty
Brookhaven National Laboratory

The Good News

- Overall, by far the smoothest DAQ run we have had… the last couple of weeks of both d-Au and p-p we had high uptime, DAQ Operators were able to solve most problems
- We took a lot of data at rates near 1 kHz
- Chris enforced the build protocol, and we did not suffer code meltdown
- Initialization was a little faster, and you didn't have to do it as often; most of the servers, including the SEB's and ATP's, could stay up for at least a day
- Event length kept to 100 kbyte even with MVD, MUTR.N, and more TEC
- The DCM's caused essentially no downtime (ok, Chi, I guess the token passing isn't so bad after all)
- How about that ARCNET server?

The Promises

So how did Run 3 compare to what was promised in DAQFEST 2002?

Summary from January 3-4, 2002 DAQFEST:

http://www.phenix.bnl.gov/~phoncs/daqfest_2002_01/

1.The performance goals are to push events from the DCM through the Event Builder to the ATP's at the highest speed consistent with the FEM's and multiplexed DCM's, which should be near 8 kHz. The goal of the data logging is to log data at 100 Mbyte/sec or more. 100 Mbyte/sec logging, but only 1.2 kHz event rate.
2.Many fundamental software components will be upgraded at the end of the run, as soon as possible. However, phoncs0 will remain intact with the final version of online software frozen as much as possible. That was done, but we never looked back.
3.Right after the run with the detector still intact, several DAQ issues could be investigated if the power can be kept on:
- the convert and transmission times Not done.
- throughput of the event builder Not done.

4.Reliability of some of the FEM's should be investigated, particularly the MVD, TEC, and initial download of the EMC. Low Voltage Power Supply reliability should be studied and improved if possible. Much better; few FEM failures after MUTR.N repair.
5.Some FEM's require further development.
- The five event buffer should be tested and fixed where necessary (PC) Still not quite done, but close
- EMC short format should be made to work and zero suppressed Not finished either, but works in majority of FEM's
- The FEM CONVERT and ENDAT times should be optimized for all systems Not done.
- The RICH FEM ARCNET complete initialization should speed up Not done; somehow not an issue.

6.Solaris, CORBA, and VxWorks will be upgraded to current versions. This will require many changes to running code to accomodate. Done, very successful.
7.It may be feasible to offload from Solaris to Linux run control and its servers, however, the availability and stability of CORBA services must be investigated as a first step. Big win! Much better.
8.Diagnostics and messaging should be improved to allow more rapid diagnosis of problems. Not much better, but wasn't a big problem.
9.It is desirable to speed up the startup and initialization of the entire DAQ as much as possible. If the upgrades to CORBA don't do this, the sources of the slowups should be investigated. Lots done, but overall not much faster, partly due to ATM configuration.
10.The EvBServer will be moved to the NT side in an attempt to speed up and parallelize the initialization of the Event Builder components, unless other upgrades obviate the need for this or make clear the origin of the slowness. Not done at all, approach abandoned.
11.Additional DCM's and SEB's with JSEB's will be needed for 4 MUTR.N granulettes and 1 MUID.N for two new granules. Done, including FCAL (not envisioned at the time).
12.JSEB problems identified in this run will be studied: the "first event" problem, the need for padding at the end of many granules, the counters for multiple events per buffer, and the use of interrupts. Padding still needed, despite thinking it was fixed; other problems fixed.
13.DSP code in the DCM's will be investigated to determine whether events can be transmitted faster than the presently measured 1.8 kHz; this may require re-coding critical code in assembler. BB measured at over 5.4 kHz.
14.The Event Builder will be upgraded and expanded with additional SEB's, ATP's. It is expected that the ATM switch will be replaced by a Gigabit Ethernet switch if studies of a Gigabit switch confirm that the performance will be improved. Used in tests, not used in any serious data taking.
15.There will be alterations to the present network architecture to incorporate several switches managed by us. Yes, sort of.
16.There are several areas of R&D which may not be ready for deployment in Run 3 but which will be pursued:
- an upgraded Partition Module which can create trigger primitives Not tested
- ways for switching clocks which are less disruptive than what is done presently Tom Kerner left the lab.
- compression of data in the ATP's or buffer boxes, among others Not tried

2002.01 vs 2003.06

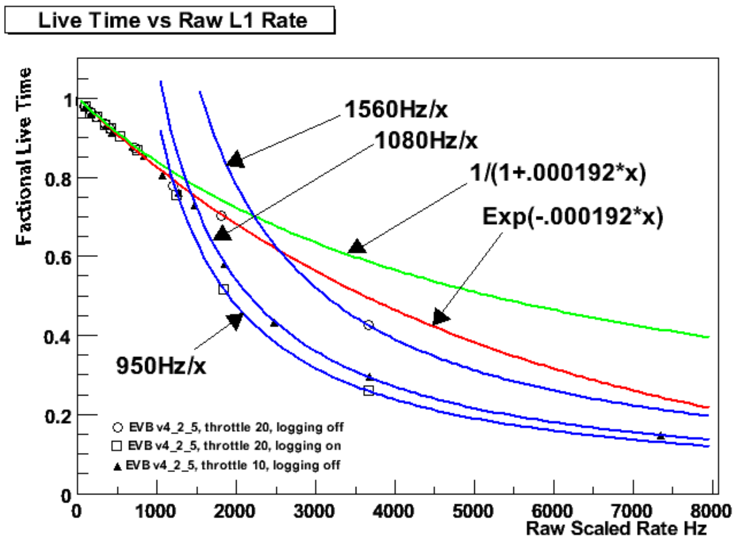
- Mostly delivered on promised improvements, which were for the most part envisioned correctly a year in advance
- The major deficiencies:
  - Multievent buffering in the FEM's is closer to being usable, but not there
  - Data doesn't get through the event builder any faster than it did at the end of 2001
  - Level 2 not in routine use
  - The data logger can't log data at full disk speed

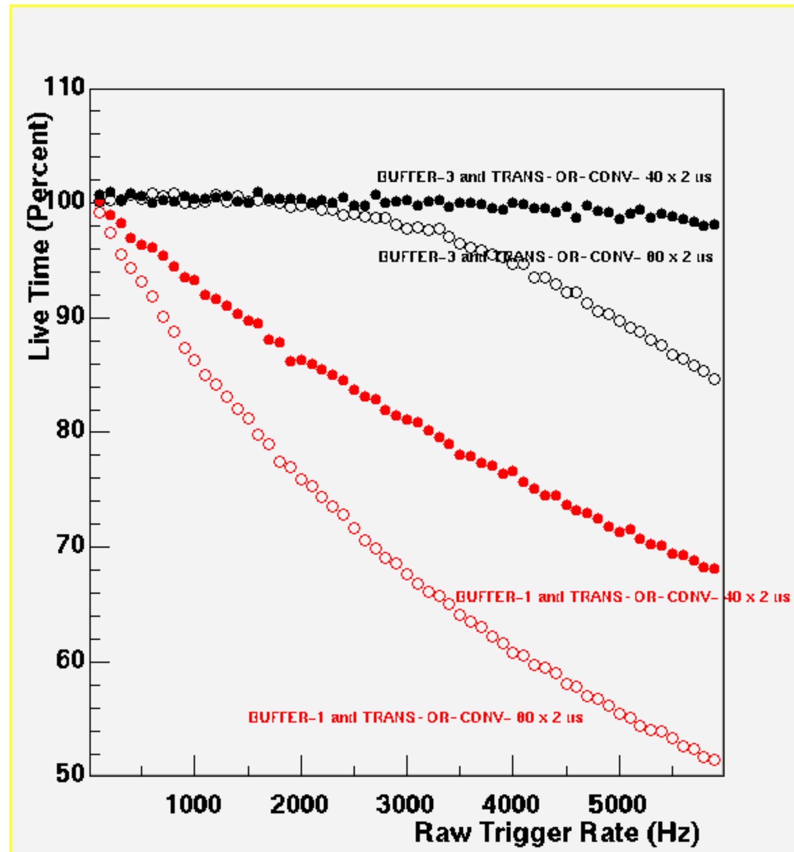
EvB Performance

From http://www.phenix.bnl.gov/~phoncs/daqfest_2002_01/cole_daq2002.ppt

- Current Status
  - Data rate
    - Have achieved rates of > 150 Mbyte/s.
    - No throughput limitations observed so far.
    - Have x3 headroom w/ existing hardware.
  - Event rate
    - Currently limited to ~ 1.3 kHz for full system.
    - Individual SEBs can run as fast as 2.3 kHz.
      - e.g. GL1 SEB On 933 MHz machines
      - More typically: 1.8 kHz for "larger" granules
    - Evidence that event rate depends on data volume (issue w/ PCI performance ?)

Steve Adler meets Jamie Nagle



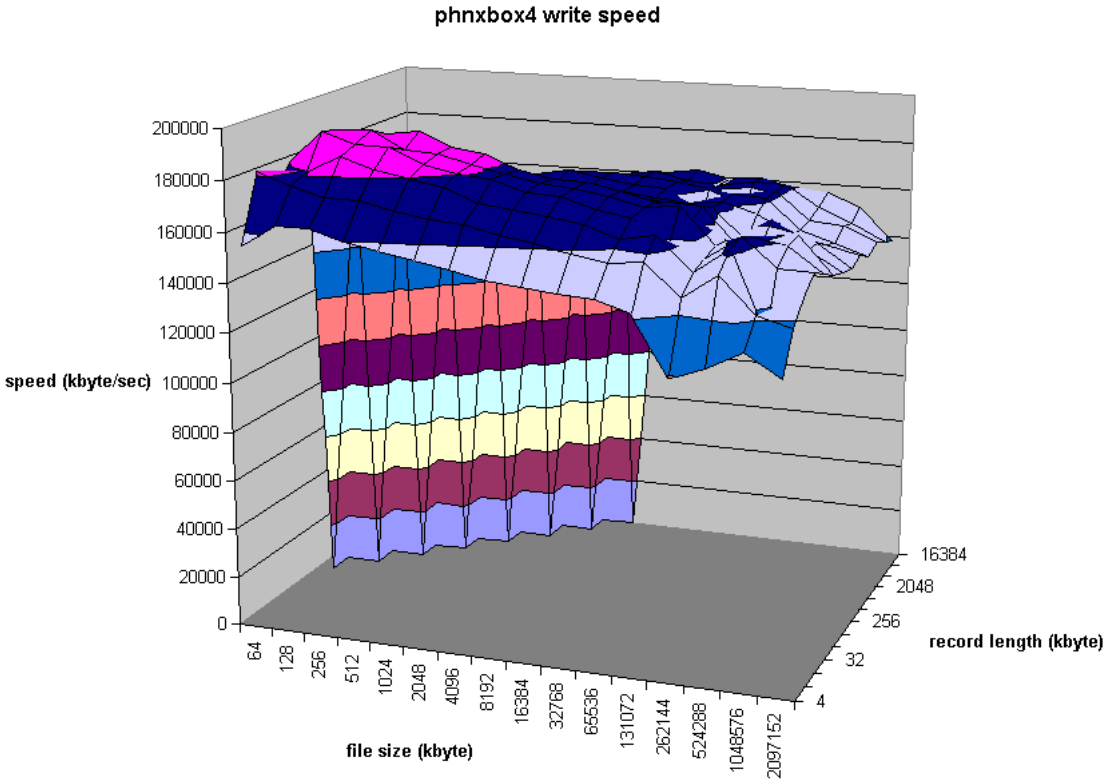


Multievent Buffering

- We have Miljko through the summer if necessary (except mid-June through mid-July, when he's on vacation in Bosnia).
- He may be able to clean up the remaining problems before he leaves, or he may have to dig into the "converter" FPGA, too.
- He has been making pretty fast progress, although it would be faster if he had a better test setup (like the West Carriage on the bench).
- Paul is working on the EMCAL to understand why some FEM's don't work with the new FPGA code, which gives us some variation of short format and a more mature multievent buffering
- Chi ironed out difficulties with BB and TOF
- Other systems appear to be ok… we need to look for trouble, though, before the next run. Chi?

Data Logging Speed

- Looks like the disk should be able to write at above 140 Mbyte/sec….



EvB Development

A number of things need investigating and coding:

- **We've got to fix the SEB code so that it uses the JSEB the way the hardware was designed to be used… the way it is, it could be the major bottleneck in the whole system**
- I need to figure out why we still need padding
- The transfer of data from the ATP's to buffer boxes has to be optimized in tandem with the logger; it must accommodate ATP's falling off and other routine disasters gracefully
- The Gigabit switch has to be integrated and managed; we need to see some (all?) granules pushing data to ATP's at "full" Gigabit speed
- The Level 2 needs to run for 8 hours on pedestals (then the algorithms must be verified by the Code Police and tested on at least 1M events)
- The flow must go! Events of all sizes and rates must be accomodated gracefully.
- All machines should run the same OS

What Can We Do to Help?

I haven't found anyone outside Nevis who wants to work on the Microsoft Windows part of the code, so Brian's going to have to provide manpower to work on that. Maybe we can help in other ways… these are tasks that people have told me they would work on if it would help and they can make progress:

- **I can tell you how to use the JSEB efficiently… I might even be able to debug the readout if pointed at it and were able to build it and could run the SEB standalone.**
- Martin (maybe me, too) could take on the network management and monitoring of the of the Gigabit switch
- If there are CORBA calls in the EvB that could be made into a diagnostic monitor, Ed could help create a client that displays it or logs it
- Martin could work with someone to optimize on both sides the data logging, and try compression

Proposal for further development

Those things are all well and good, but we would probably be more effective if we debugged a "toy" event builder in our native language to figure out the problems ourselves. To this end, we would need:

- Buy the Linux version of Jungo ($2k), test it at speed (probably provides useful diagnostic capability anyway, and future flexibility)
- Take a couple of granules (GL1, BB?) and create SEB code that tests the throughput; add some extra cables so we can move back and forth easily
- Take a couple VA machines and test ATP-buffer box networking and logging

Where to from here?

- We need to make progress on the event handling ability
- We need to make it soon in case there are gotchas that will take some time to be corrected
- We have everything at hand that we need to understand the rate limitations
- We need to agree on a course of action in this meeting, at least until September
- We need to reconvene in September to assess where we've been and where we're headed